

Talking toxicogenomics and global database

By Ernie Hood

The daunting challenges and astonishing opportunities presented by big data are currently attracting attention in many fields, including toxicogenomics, where technology is contributing to rapid expansion in the number and size of databases, most of which do not presently interact effectively. That premise spawned an international meeting at NIEHS June 26-27 titled Workshop on Identifying Opportunities for Global Integration of Toxicogenomics Databases.

More than 30 scientists from around the world, involved with toxicogenomics databases, attended the workshop, which was organized over the past year by Allen Dearry, Ph.D., and Rebecca Boyles, of the [NIEHS Office of Scientific Information Management](#); Jennifer Fostel, Ph.D., NTP scientific administrator of the [Chemical Effects in Biological Systems \(CEBS\)](#) database; and their colleagues Jos Kleinjans, Ph.D., and Diana Hendrickx, Ph.D., from Maastricht University in the Netherlands, who are involved with European Union (EU) toxicogenomics database initiatives.

An array of toxicogenomics databases

The first day of the workshop was devoted to presentations about the many toxicogenomics databases around the world, all of which are somewhat different in their contents and approaches. "It was good for everybody to see the scope of all of the various toxicogenomic databases, as people got some familiarity with the universe of databases that are out there," said Dearry. "And then the next step is to try to see what we can do to start to enable cross talk between them."

During the second day of the gathering, the group engaged in discussions designed to yield concrete action plans. Participants identified three types of problems in need of attention - database communication, such as metadata consensus and definition of standards; sustainability, such as funding continuity and data preservation; and the need for education in the field to ensure an adequate and trained future workforce.

First things first

Attendees quickly recognized that, in order to effectively and efficiently tackle such a complex undertaking, solid information about the wants, needs, and goals involved will be necessary prior to moving forward. Boyles posed provocative initial questions to the group about what they expect to gain by interoperability in this field. "This is a lot of time, effort, and, therefore, money, so you'd better be sure that what we're pursuing is going to add some value to the field."

With those challenges in mind, step one of the group's action plan involves research. "The action plan really has some concrete steps in terms of developing our own inventory of what the toxicogenomics databases are, going out and searching for some use cases about toxicogenomics research and how these databases can be useful in that, and then using all of that information to develop a white paper to try to explain what this integration can bring about and why it would be useful to the larger research community," Boyles explained.

The goal is to have that research accomplished by mid-September, in time for the Research Data Alliance (RDA) [Second Plenary Meeting](#),

(<https://rd-alliance.org/future-events>)

Building Global Partnerships, Sept. 16-18 at the National Academy of Sciences in Washington, D.C. The RDA is a consortium arrangement among organizations in the U.S., EU, and Australia that aims to accelerate and facilitate research data sharing and exchange. Much of its work is accomplished by its working groups and interest groups, so the intent would be to propose establishment of an RDA special interest group to provide a formal mechanism and framework for enhancing interoperability of toxicogenomics databases.

(Ernie Hood is a contract writer with the NIEHS Office of Communications and Public Liaison.)



"We've been a victim of our own success," Dearry told the workshop audience. "It's become easier to produce data than it is to store and analyze and disseminate those data. So how can we improve our data management and data handling?" (Photo courtesy of Steve McCaw)



Kleinjans noted the accelerating progress in the field, with rapidly expanding data sets. "What is relevant for all of us is that toxicogenomics, as a technological approach, has not yet stabilized. Sequencing technologies will replace the microarray-based approach we have applied over the last decade or so, and will create an even bigger, a tremendously bigger, data set, which will enable us to go even deeper into the biology." (Photo courtesy of Steve McCaw)



Weida Tong, Ph.D., director of the FDA Center for Toxicoinformatics, briefed workshop participants on the [MicroArray Quality Control](http://www.fda.gov/ScienceResearch/BioinformaticsTools/MicroarrayQualityControlProject/default.htm) (<http://www.fda.gov/ScienceResearch/BioinformaticsTools/MicroarrayQualityControlProject/default.htm>) project, the FDA-led, community-wide consortium effort to assess technical performance and application of genomics technologies. (Photo courtesy of Steve McCaw)



Alex Merrick, Ph.D., leader of the Molecular Toxicology and Informatics Group in the NTP Biomolecular Screening Branch, was one of more than 30 workshop participants, who brought a wide range of expertise to the meeting. Merrick joined NTP molecular toxicologist Scott Auerbach, Ph.D., who described the [DrugMatrix](https://ntp.niehs.nih.gov/drugmatrix/index.html) (<https://ntp.niehs.nih.gov/drugmatrix/index.html>) database, a large toxicogenomics reference resource. (Photo courtesy of Steve McCaw)



Workshop attendees, such as NTP biologist Laura Hall, who took careful notes about the presentations, learned about the nuts and bolts of many different toxicogenomics databases, including large repositories in the U.S., EU, and Japan. (Photo courtesy of Steve McCaw)



Stephen Edwards, Ph.D., a systems biologist with the EPA National Health and Environmental Effects Research Laboratory, presented an overview of "Toxicogenomics data at the U.S. EPA." (Photo courtesy of Steve McCaw)

The Environmental Factor is produced monthly by the [National Institute of Environmental Health Sciences \(NIEHS\)](http://www.niehs.nih.gov/)

(<http://www.niehs.nih.gov/>)

, Office of Communications and Public Liaison. The content is not copyrighted, and it can be reprinted without permission. If you use parts of Environmental Factor in your publication, we ask that you provide us with a copy for our records. We welcome your [comments and suggestions](#). (bruskec@niehs.nih.gov)

This page URL: NIEHS website: <http://www.niehs.nih.gov/>
Email the Web Manager at webmanager@niehs.nih.gov